BIHOMOGENEOUS FORMS IN MANY VARIABLES

DAMARIS SCHINDLER

Abstract. We count integer points on bihomogeneous varieties using the Hardy-Littlewood method. The main novelty lies in using the structure of bihomogeneous equations to obtain asymptotics in generically fewer variables than would be necessary in using the standard approach for homogeneous varieties. Also, we consider counting functions where not all the variables have to lie in intervals of the same size, which arises as a natural question in the setting of bihomogeneous varieties.

1. Introduction

An important issue in the study of diophantine equations is to determine the density of integer points on algebraic varieties. In this setting the circle method is a powerful instrument, with which for example Birch [1] and Schmidt [7] obtained results in great generality. So far, most literature is concerned with counting integer points in boxes which are dilated by a large real number. In this case all the variables lie in intervals of comparable length. In this paper we study systems of bihomogeneous equations where it is natural to ask for similar asymptotic formulas while allowing different sizes for the variables involved. Furthermore, we use the structure of bihomogeneous equations to obtain results on the number of integer points on these varieties, using in generic cases fewer variables than needed in Birch's work [1].

First we need to introduce some notation. Let n_1, n_2 and R be positive integers. We use the vector notation $\mathbf{x} = (x_1, \ldots, x_{n_1})$ and $\mathbf{y} = (y_1, \ldots, y_{n_2})$. We call a polynomial $F(\mathbf{x}; \mathbf{y}) \in \mathbb{Z}[\mathbf{x}, \mathbf{y}]$ a bihomogneneous form of bidegree (d_1, d_2) if

$$F(\lambda\mathbf{x}; \mu\mathbf{y}) = \lambda^{d_1}\mu^{d_2}F(\mathbf{x}; \mathbf{y}),$$

for all $\lambda, \mu \in \mathbb{C}$ and all vectors $\mathbf{x}, \mathbf{y}$. In the following we consider a system of bihomogeneous forms $F_i(\mathbf{x}, \mathbf{y}) \in \mathbb{Z}[\mathbf{x}, \mathbf{y}]$, for $1 \leq i \leq R$. We are interested in the number of solutions to the system of equations

$$F_i(\mathbf{x}; \mathbf{y}) = 0, \tag{1.1}$$

for $1 \leq i \leq R$, where we seek integer solutions in certain boxes. Thus, let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two boxes of side length at most 1 in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, and let P_1 and P_2 be large real numbers. We write $P_1\mathcal{B}_1$ for the set of $\mathbf{x} \in \mathbb{R}^{n_1}$ such that $P_1^{-1}\mathbf{x} \in \mathcal{B}_1$, and $P_2\mathcal{B}_2$ analogously. Then we define $N(P_1, P_2)$ to be the number

2010 *Mathematics Subject Classification.* 11D45 (11D72, 11P55).

Key words and phrases. bihomogeneous equations, Hardy-Littlewood method.

of integer solutions to the system of equations (1.1) with

$$\mathbf{x} \in P_1\mathcal{B}_1 \text{ and } \mathbf{y} \in P_2\mathcal{B}_2.$$

Furthermore, we introduce the affine variety V_1^* in $\mathbb{A}_{\mathbb{C}}^{n_1+n_2}$ given by

$$\operatorname{rank}\left(\frac{\partial F_i}{\partial x_j}\right)_{\substack{1\leq i\leq R\\ 1\leq j\leq n_1}} < R. \tag{1.2}$$

Similarly we define V_2^* to be the affine variety in $\mathbb{A}_{\mathbb{C}}^{n_1+n_2}$ given by

$$\operatorname{rank}\left(\frac{\partial F_i}{\partial y_j}\right)_{\substack{1\leq i\leq R\\ 1\leq j\leq n_2}} < R. \tag{1.3}$$

Our main result is an asymptotic formula for $N(P_1, P_2)$, which we can establish as soon as the codimensions of V_1^* and V_2^* are sufficiently large in terms of the number of equations, the bidegree of the polynomials and the logarithmic ratio between the two parameters P_1 and P_2.

Theorem 1.1. *Let P_1 and P_2 be two large real numbers, and define $b = \frac{\log P_1}{\log P_2}$. Assume that $b \geq 1$. Furthermore, for all $1 \leq i \leq R$, assume that the polynomials F_i have bidegree (d_1, d_2). Let $n_1, n_2 > R$ and V_1^* and V_2^* be the varieties given by equations (1.2) and (1.3). Assume that*

$$n_1 + n_2 - \dim V_i^* > 2^{d_1+d_2-2} \max\{R(R+1)(d_1 + d_2 - 1), R(bd_1 + d_2)\},$$

for $i = 1, 2$. Then we have the asymptotic formula

$$N(P_1, P_2) = \sigma P_1^{n_1-Rd_1} P_2^{n_2-Rd_2} + O(P_1^{n_1-Rd_1-\varepsilon} P_2^{n_2-Rd_2}),$$

for some real σ and $\varepsilon > 0$. As usual, σ is the product of a singular series $\mathfrak{S}$ and a singular integral J which are given in equations (5.6) and (5.7). Furthermore, the constant σ is positive if

i) the $F_i(\mathbf{x}; \mathbf{y})$ have a common non-singular p-adic zero for all p,

ii) and if the $F_i(\mathbf{x}; \mathbf{y})$ have a non-singular real zero in the box $\mathcal{B}_1 \times \mathcal{B}_2$ and $\dim V(0) = n_1 + n_2 - R$, where $V(0)$ is the affine variety given by the system of equations (1.1).

We note that in our result the number of variables n_1 and n_2 depends on the parameter b. However, this condition can be omitted if

$$(R + 1)(d_1 + d_2 - 1) \geq (bd_1 + d_2).$$

There are few examples in the literature where the number of integer points on bihomogeneous varieties is studied. Robbiani ([5]) and Spencer ([8]) treat bilinear varieties, and Van Valckenborgh ([9]) provides some results on bihomogeneous equations of bidegree $(2, 3)$. However, Van Valckenborgh only considers a diagonal situation, whereas we are interested in a general set-up.

In our work we largely follow Birch's paper [1]. However, we have to take care of the different sizes of our boxes and their growth. The main difference to Birch's work is in the form of Weyl's inequality we use. When Birch works with forms of total degree d he differentiates them $d - 1$ times via Weyl-differencing to obtain linear exponential sums. We apply that differencing

process separately with respect to the variables $\mathbf{x}$ and $\mathbf{y}$, such that we only have to use this process d_1-1 times for the variables $\mathbf{x}$ and d_2-1 times for the variables $\mathbf{y}$. In total we therefore only need d_1+d_2-2 differencing steps. This approach was first mentioned to us by Prof. T. D. Wooley. One condition in Birch's theorem is that the total number of variables $\tilde{n}$ satisfies

$$\tilde{n} - \dim V^* > R(R+1)(d-1)2^{d-1},$$

which is essentially determined by the form of Weyl's lemma, which he uses. We obtain a similar condition for $d = d_1 + d_2$, however we can replace the factor 2^{d-1} by 2^{d-2}.

On the other hand, in our condition the quantities $\dim V_1^*$ and $\dim V_2^*$ appear instead of the dimension of V^*, which is the variety given by

$$\operatorname{rank}\left(\frac{\partial F_i}{\partial z_j}\right) < R,$$

where z_j run through all variables $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$. We clearly have $V^* \subset V_i^*$ and thus $\dim V^* \leq \dim V_i^*$, for $i = 1, 2$. However, we note that the singular locus of a bihomogeneous variety is rather large, as soon as not both d_1 and d_2 equal 1. If we assume for example $d_1 > 1$, then we see that V^* contains a linear subspace of dimension n_2, when we set $\mathbf{x} = 0$. The same holds of course for V_1^* and V_2^*. We assume for the moment that we have $n = n_1 = n_2$ and that d_1 or d_2 is larger than 1. Then we claim that in a generic situation we have

$$n = \dim V^* = \dim V_1^* = \dim V_2^*. \tag{1.4}$$

Since each of the loci has dimension at least n, and $V^* \subset V_1^*$, it suffices by symmetry to show that $\dim V_1^* = n$ in the generic situation.

To justify this claim, we note that for fixed bidegree (d_1, d_2) with $d_1, d_2 \geq 1$ there are

$$m = \binom{n+d_1-1}{n-1}\binom{n+d_2-1}{n-1}$$

monomials of bidegree (d_1, d_2) in $(\mathbf{x}; \mathbf{y})$. We fix an order of them and associate to each $\mathbf{a} \in \mathbb{A}_{\mathbb{Q}}^m$ a bihomogeneous form $F_{\mathbf{a}}(\mathbf{x}; \mathbf{y})$. We write $\nabla_{\mathbf{x}} F$ for the gradient of a bihomogeneous form $F(\mathbf{x}; \mathbf{y})$ with respect to the variables $\mathbf{x}$. For $\mathbf{a} \in \mathbb{P}_{\mathbb{Q}}^{m-1}$ we set

$$X_{1,\mathbf{a}} = \{(\mathbf{x}; \mathbf{y}) \in \mathbb{P}_{\mathbb{Q}}^{n-1} \times \mathbb{P}_{\mathbb{Q}}^{n-1} : \nabla_{\mathbf{x}} F_{\mathbf{a}}(\mathbf{x}; \mathbf{y}) = 0\}.$$

Furthermore, we consider the projective variety

$$\mathcal{V} = \{(\mathbf{a}; \mathbf{x}; \mathbf{y}) \in \mathbb{P}_{\mathbb{Q}}^{m-1} \times \mathbb{P}_{\mathbb{Q}}^{n-1} \times \mathbb{P}_{\mathbb{Q}}^{n-1} : \nabla_{\mathbf{x}} F_{\mathbf{a}}(\mathbf{x}; \mathbf{y}) = 0\},$$

and the projection to the first factor $\pi : \mathcal{V} \to \mathbb{P}_{\mathbb{Q}}^{m-1}$. Define the function

$$\lambda(\mathbf{a}) = \dim(\pi^{-1}(\mathbf{a})) = \dim X_{1,\mathbf{a}},$$

for $\mathbf{a} \in \mathbb{P}_{\mathbb{Q}}^{m-1}$. Then Corollary 11.13 of [4] shows that λ is an upper semi-continuous function on $\pi(\mathcal{V})$ in the Zariski-topology of $\pi(\mathcal{V})$, which is itself a closed subset of $\mathbb{P}_{\mathbb{Q}}^{m-1}$ by Theorem 3.13 of [4]. Hence the set

$$Y = \{\mathbf{a} \in \mathbb{P}_{\mathbb{Q}}^{m-1} : \lambda(\mathbf{a}) \geq n-1\}$$

is closed in $\pi(\mathcal{V})$ and hence in $\mathbb{P}_{\mathbb{Q}}^{m-1}$. We claim that $Y \neq \mathbb{P}_{\mathbb{Q}}^{m-1}$. For this we consider the vector $\mathbf{b} \in \mathbb{A}_{\mathbb{Q}}^m \setminus \{0\}$ such that

$$F_{\mathbf{b}}(\mathbf{x};\mathbf{y}) = x_1^{d_1} y_1^{d_2} + \ldots + x_n^{d_1} y_n^{d_2}.$$

Then $X_{1,\mathbf{b}}$ is given by $x_i y_i = 0$ for $1 \leq i \leq n$ if $d_1 \geq 2$, and empty if $d_1 = 1$. In any case, we have $\dim X_{1,\mathbf{b}} \leq n-2$. Therefore the set

$$\{\mathbf{a} \in \mathbb{P}_{\mathbb{Q}}^{m-1} : \dim X_{1,\mathbf{a}} \leq n-2\}$$

is open and non-empty in $\mathbb{P}_{\mathbb{Q}}^{m-1}$, and so $\dim V_1^* = n$ in the generic case..

Another novelty in this work is the way we use of the geometry of numbers in the treatment of our exponential sums. Birch in his paper [1] uses Lemma 12.6 from [3], which is a standard argument at this step. However, this lemma can only be applied if the involved matrices are symmetric, which is not the case in our situation. Our Lemma 3.1 provides a form of generalising that lemma from Davenport to general matrices.

We note that a system of bihomogeneous polynomials $F_i(\mathbf{x};\mathbf{y})$ defines a variety in biprojective space $\mathbb{P}^{n_1-1} \times \mathbb{P}^{n_2-1}$. Hence, in the context of the Manin conjectures, it is natural to count rational points on this variety with respect to the anticanonical height function in biprojective space. Our Thoerem 1.1 is a first step in this direction and will be used to accomplish this goal in forthcoming work of the author. We note that it will turn out to be important that we can establish asymptotic formulas for $N(P_1, P_2)$ for parameters P_1 and P_2 which are not necessarily of the same size.

In the following $\boldsymbol{\alpha}$ is some vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_R) \in \mathbb{R}^R$, and we use the abbreviation $\boldsymbol{\alpha} \cdot \mathbf{F} := \alpha_1 F_1 + \ldots + \alpha_R F_R$. Furthermore, we frequently use summations over integer vectors $\mathbf{x}$ and $\mathbf{y}$, such that sums of the type $\sum_{\mathbf{x} \in P_1 \mathcal{B}_1}$ are to be understood as sums $\sum_{\mathbf{x} \in P_1 \mathcal{B}_1 \cap \mathbb{Z}^{n_1}}$. For a real number x we write $\|x\| = \min_{z \in \mathbb{Z}} |x - z|$ for the distance to the nearest integer. As usual, we write $e(x)$ for $e^{2\pi i x}$.

The structure of this paper is as follows. After introducing some notation in section 2, we perform a Weyl-differencing process in section 3. In section 4 we are concerned with the lemma from the geometry of numbers mentioned above. This is used in section 5 to deduce a form of Weyl's inequality. In section 6 we set up the circle method, reduce the problem to a major arc situation and treat the singular series and integral. The proof of Theorem 1.1 is finished in the final section.

Acknowledgements. During part of the work on this paper the author was supportet by a DAAD scholarship. Furthermore, the author would like to thank Prof. T. D. Wooley for suggesting this area of research.

2. Exponential sums

We start in defining the exponential sum

$$S(\boldsymbol{\alpha}) = \sum_{\mathbf{x} \in P_1 \mathcal{B}_1} \sum_{\mathbf{y} \in P_2 \mathcal{B}_2} e(\boldsymbol{\alpha} \cdot \mathbf{F}(\mathbf{x};\mathbf{y})),$$

for some $\boldsymbol{\alpha} \in \mathbb{R}^R$. One goal of this section is to perform (d_1-1) times a Weyl-differencing process with respect to the variables $\mathbf{x}$ and (d_2-1) times the same differencing process with respect to $\mathbf{y}$. For this we write each bihomogeneous form F_i as

$$F_i(\mathbf{x};\mathbf{y}) = \sum_{\mathbf{j}=1}^{n_1}\sum_{\mathbf{k}=1}^{n_2} F^{(i)}_{j_1,\ldots,j_{d_1};k_1,\ldots,k_{d_2}} x_{j_1}\ldots x_{j_{d_1}} y_{k_1}\ldots y_{k_{d_2}},$$

with the $F^{(i)}_{j_1,\ldots,j_{d_1};k_1,\ldots,k_{d_2}}$ symmetric in $(j_1,\ldots,j_{d_1})$ and $(k_1,\ldots,k_{d_2})$. Here the summations are over $j_1,\ldots,j_{d_1}$ from 1 to n_1, and $k_1,\ldots,k_{d_2}$ from 1 to n_2, and we write $\mathbf{j}$ and $\mathbf{k}$ for $(j_1,\ldots,j_{d_1})$ and $(k_1,\ldots,k_{d_2})$. Without loss of generality we can assume the $F^{(i)}_{\mathbf{j};\mathbf{k}}$ to be integers (otherwise multiply with some suitable constant).

Let $d_2>1$. We start our differencing process in applying Hölder's inequality to obtain

$$|S(\boldsymbol{\alpha})|^{2^{d_2-1}} \ll P_1^{n_1(2^{d_2-1}-1)} \sum_{\mathbf{x}\in P_1\mathcal{B}_1} |S_{\mathbf{x}}(\boldsymbol{\alpha})|^{2^{d_2-1}}, \tag{2.1}$$

with the exponential sum

$$S_{\mathbf{x}}(\boldsymbol{\alpha}) = \sum_{\mathbf{y}\in P_2\mathcal{B}_2} e(\boldsymbol{\alpha}\cdot\mathbf{F}(\mathbf{x};\mathbf{y})).$$

Next we use a form of Weyl's inequality as in Lemma 11.1 in [7] to bound $|S_{\mathbf{x}}(\boldsymbol{\alpha})|^{2^{d_2-1}}$. For this we need to introduce some notation. Let $\mathcal{U} = P_2\mathcal{B}_2$, write $\mathcal{U}^D = \mathcal{U}-\mathcal{U}$ for the difference set and define

$$\mathcal{U}(\mathbf{y}^{(1)},\ldots,\mathbf{y}^{(t)}) = \cap_{\varepsilon_1=0}^{1}\ldots\cap_{\varepsilon_t=0}^{1}(\mathcal{U}-\varepsilon_1\mathbf{y}^{(1)}-\ldots-\varepsilon_t\mathbf{y}^{(t)}).$$

Following the notation of [7], we define the polynomial $\mathcal{F}(\mathbf{y}) = \boldsymbol{\alpha}\cdot\mathbf{F}(\mathbf{x};\mathbf{y})$. Furthermore we set

$$\mathcal{F}_d(\mathbf{y}_1,\ldots,\mathbf{y}_d) = \sum_{\varepsilon_1=0}^{1}\ldots\sum_{\varepsilon_d=0}^{1}(-1)^{\varepsilon_1+\ldots+\varepsilon_d}\mathcal{F}(\varepsilon_1\mathbf{y}_1+\ldots+\varepsilon_d\mathbf{y}_d),$$

and $\mathcal{F}_0=0$ identically.

In our estimate for $|S_{\mathbf{x}}(\boldsymbol{\alpha})|^{2^{d_2-1}}$ we want to avoid absolute values in the resulting bound such that we directly consider equation 11.2 in [7]. This delivers the estimate

$$|S_{\mathbf{x}}(\boldsymbol{\alpha})|^{2^{d_2-1}} \ll |\mathcal{U}^D|^{2^{d_2-1}-d_2} \sum_{\mathbf{y}^{(1)}\in\mathcal{U}^D}\ldots \sum_{\mathbf{y}^{(d_2-2)}\in\mathcal{U}^D}\left|\sum_{\mathbf{y}^{(d_2-1)}\in\mathcal{U}(\mathbf{y}^{(1)},\ldots\mathbf{y}^{(d_2-2)})} e(\mathcal{F}_{d_2-1}(\mathbf{y}^{(1)},\ldots,\mathbf{y}^{(d_2-1)}))\right|^2,$$

We note that all the summation regions for the $\mathbf{y}^{(j)}$ are boxes, since $P_2\mathcal{B}_2$ is a box and intersections and differences of boxes are again boxes. As in the proof

of Lemma 11.1 in [7] we consider two elements $\mathbf{z}, \mathbf{z}' \in \mathcal{U}(\mathbf{y}^{(1)}, \dots \mathbf{y}^{(d_2-2)})$ and note that

$$\begin{aligned} &\mathcal{F}_{d_2-1}(\mathbf{y}^{(1)}, \dots, \mathbf{z}) - \mathcal{F}_{d_2-1}(\mathbf{y}^{(1)}, \dots, \mathbf{z}') \\ =&\mathcal{F}_{d_2-1}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-2)}, \mathbf{y}^{(d_2)}) - \mathcal{F}_{d_2-1}(\mathbf{y}^{(1)}, \dots \mathbf{y}^{(d_2-2)}, \mathbf{y}^{(d_2-1)} + \mathbf{y}^{(d_2)}) \\ =&\mathcal{F}_{d_2}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-1)}, \mathbf{y}^{(d_2)}) - \mathcal{F}_{d_2-1}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-2)}, \mathbf{y}^{(d_2-1)}), \end{aligned}$$

for some $\mathbf{y}^{(d_2-1)} \in \mathcal{U}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-2)})^D$ and $\mathbf{y}^{(d_2)} \in \mathcal{U}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-1)})$. Thus, we obtain the bound

$$\begin{aligned} |S_{\mathbf{x}}(\boldsymbol{\alpha})|^{2^{d_2-1}} \ll P_2^{n_2(2^{d_2-1}-d_2)} &\sum_{\mathbf{y}^{(1)} \in \mathcal{U}^D} \dots \sum_{\mathbf{y}^{(d_2-2)} \in \mathcal{U}^D} \sum_{\mathbf{y}^{(d_2-1)} \in \mathcal{U}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-2)})^D} \\ &\sum_{\mathbf{y}^{(d_2)} \in \mathcal{U}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-1)})} e(\mathcal{F}_{d_2}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2)}) - \mathcal{F}_{d_2-1}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-1)})). \end{aligned}$$

By Lemma 11.4 of Schmidt's work [7] the polynomial $\mathcal{F}_{d_2}$ is just the multilinear form associated to $\mathcal{F}$. In our case we have

$$\begin{aligned} &\mathcal{F}_{d_2}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2)}) - \mathcal{F}_{d_2-1}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-1)}) \\ =&\sum_{i=1}^{R} \alpha_i \sum_{\mathbf{j}} \sum_{\mathbf{k}} F_{\mathbf{j},\mathbf{k}}^{(i)} x_{j_1} \dots x_{j_{d_1}} h_{\mathbf{k}}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2)}), \end{aligned}$$

with

$$h_{\mathbf{k}}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2)}) = d_2! y_{k_1}^{(1)} \dots y_{k_{d_2}}^{(d_2)} + \tilde{h}_{\mathbf{k}}(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2-1)}),$$

where $\tilde{h}_{\mathbf{k}}$ are some homogeneous polynomials of degree d_2 independent of $\mathbf{y}^{(d_2)}$.

We come back to estimating $\sum_{\mathbf{x} \in P_1 \mathcal{B}_1} |S_{\mathbf{x}}(\boldsymbol{\alpha})|^{2^{d_2-1}}$. Set $\tilde{d} = d_1 + d_2 - 2$. We write and $\widetilde{\mathbf{y}} = (\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d_2)})$ and set

$$S_{\widetilde{\mathbf{y}}}(\boldsymbol{\alpha}) = \sum_{\mathbf{x} \in P_1 \mathcal{B}_1} e\left(\sum_i \alpha_i \sum_{\mathbf{j}} \sum_{\mathbf{k}} F_{\mathbf{j},\mathbf{k}}^{(i)} x_{j_1} \dots x_{j_{d_1}} h_k(\widetilde{\mathbf{y}}) \right).$$

In equation (2.1) we interchange the summation over $\sum_{\mathbf{x}}$ with all the summations $\sum_{\mathbf{y}^{(i)}}$ from the bound for $\sum_{\mathbf{x} \in P_1 \mathcal{B}_1} |S_{\mathbf{x}}(\boldsymbol{\alpha})|^{2^{d_2-1}}$. An application of Hölder's inequality now delivers

$$|S(\boldsymbol{\alpha})|^{2^{\tilde{d}}} \ll P_1^{n_1(2^{\tilde{d}} - 2^{d_1-1})} P_2^{n_2(2^{\tilde{d}} - d_2)} \sum_{\mathbf{y}^{(1)}} \dots \sum_{\mathbf{y}^{(d_2)}} |S_{\widetilde{\mathbf{y}}}(\boldsymbol{\alpha})|^{2^{d_1-1}}.$$

Applying the same differencing process as before to $S_{\widetilde{\mathbf{y}}}(\boldsymbol{\alpha})$ leads us to

$$|S(\boldsymbol{\alpha})|^{2^{\tilde{d}}} \ll P_1^{n_1(2^{\tilde{d}} - d_1)} P_2^{n_2(2^{\tilde{d}} - d_2)} \sum_{\mathbf{y}^{(1)}} \dots \sum_{\mathbf{y}^{(d_2)}} \sum_{\mathbf{x}^{(1)}} \dots |\sum_{\mathbf{x}^{(d_1)}} e(\gamma(\widetilde{\mathbf{x}}; \widetilde{\mathbf{y}}))|, \qquad (2.2)$$

with

$$\gamma(\widetilde{\mathbf{x}}; \widetilde{\mathbf{y}}) = \sum_i \alpha_i \sum_{\mathbf{j}} \sum_{\mathbf{k}} F_{\mathbf{j},\mathbf{k}}^{(i)} g_{\mathbf{j}}(\widetilde{\mathbf{x}}) h_{\mathbf{k}}(\widetilde{\mathbf{y}}).$$

As before we have

$$g_{\mathbf{j}}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d_1)}) = d_1! x_{j_1}^{(1)} \dots x_{j_{d_1}}^{(d_1)} + \tilde{g}_j(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d_1-1)}),$$

with some homogeneous form $\tilde{g}_{\mathbf{j}}$ of degree d_1,and all summations over $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d_1)}$ run over intervals of length at most $2P_1$. Note that equation (2.2) holds for all integers $d_1 \geq 1$ and $d_2 \geq 1$. Next we introduce the notation $\widehat{\mathbf{x}} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d_1-1)})$ and $\widehat{\mathbf{y}}$ analogously, and turn towards estimating the sum

$$\sum(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}) := \sum_{\mathbf{y}^{(d_2)}} \left| \sum_{\mathbf{x}^{(d_1)}} e(\gamma(\widetilde{\mathbf{x}}; \widetilde{\mathbf{y}})) \right|.$$

First we have

$$\left| \sum_{\mathbf{x}^{(d_1)}} e(\gamma(\widetilde{\mathbf{x}}, \widetilde{\mathbf{y}})) \right| \ll \prod_{l=1}^{n_1} \min\left(P_1, \|\tilde{\gamma}(\widehat{\mathbf{x}}, \mathbf{e}_l; \widetilde{\mathbf{y}})\|^{-1}\right),$$

where $\mathbf{e}_l$ is the lth unit vector and $\tilde{\gamma}$ is given by

$$\tilde{\gamma}(\widetilde{\mathbf{x}}; \widetilde{\mathbf{y}}) = d_1! \sum_i \alpha_i \sum_{\mathbf{j}} \sum_{\mathbf{k}} F^{(i)}_{\mathbf{j},\mathbf{k}} x^{(1)}_{j_1} \ldots x^{(d_1)}_{j_{d_1}} h_k(\widetilde{\mathbf{y}}).$$

Next we follow Davenport's analysis in [2], section 3. For some real number z we write $\{z\}$ for the fractional part, and use the notation $\mathbf{r} = (r_1, \ldots, r_n)$. For some integers $0 \leq r_l < P_1$ let $\mathcal{A}(\widehat{\mathbf{x}}; \widehat{\mathbf{y}}; \mathbf{r})$ be the set of $\mathbf{y}^{(d_2)}$ in the above summation such that

$$r_l P_1^{-1} \leq \{\tilde{\gamma}(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{y}^{(d_2)})\} < (r_l + 1) P_1^{-1},$$

for $1 \leq l \leq n_1$. Then we can estimate

$$\sum_{\mathbf{y}^{(d_2)}} \left| \sum_{\mathbf{x}^{(d_1)}} e(\gamma(\widetilde{\mathbf{x}}; \widetilde{\mathbf{y}})) \right| \ll \sum_{\mathbf{r}} A(\widehat{\mathbf{x}}; \widehat{\mathbf{y}}; \mathbf{r}) \prod_{l=1}^{n} \min\left(P_1, \max\left(\frac{P_1}{r_l}, \frac{P_1}{P_1 - r_l - 1}\right)\right),$$

where the summation is over all vectors $\mathbf{r}$ with $0 \leq r_l < P_1$ for all l, and $A(\widehat{\mathbf{x}}; \widehat{\mathbf{y}}; \mathbf{r})$ is the cardinality of the set $\mathcal{A}(\widehat{\mathbf{x}}; \widehat{\mathbf{y}}; \mathbf{r})$. Our next goal is to find a bound for $A(\widehat{\mathbf{x}}; \widehat{\mathbf{y}}; \mathbf{r})$, which is independent of $\mathbf{r}$. For this consider two vectors $\mathbf{u}$ and $\mathbf{v}$ counted by that quantity. Then we have

$$\|\tilde{\gamma}(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{u}) - \tilde{\gamma}(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{v})\| < P_1^{-1},$$

for $1 \leq l \leq n_1$. Define the multilinear form

$$\Gamma(\widetilde{\mathbf{x}}; \widetilde{\mathbf{y}}) = d_1! d_2! \sum_i \alpha_i \sum_{\mathbf{j}} \sum_{\mathbf{k}} F^{(i)}_{\mathbf{j},\mathbf{k}} x^{(1)}_{j_1} \ldots x^{(d_1)}_{j_{d_1}} y^{(1)}_{k_1} \ldots y^{(d_2)}_{k_{d_2}},$$

and let $N(\widehat{\mathbf{x}}; \widehat{\mathbf{y}})$ be the number of integer vectors $\mathbf{y} \in (-P_2, P_2)^{n_2}$ such that

$$\|\Gamma(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{y})\| < P_1^{-1},$$

for all $1 \leq l \leq n_1$. Observe that

$$\tilde{\gamma}(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{u}) - \tilde{\gamma}(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{v}) = \Gamma(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{u} - \mathbf{v}).$$

Thus, we have

$$A(\widehat{\mathbf{x}}; \widehat{\mathbf{y}}; \mathbf{r}) \leq N(\widehat{\mathbf{x}}; \widehat{\mathbf{y}}),$$

for all $\mathbf{r}$ under consideration. This gives us finally the bound

$$\sum_{\mathbf{y}^{(d_2)}}\left|\sum_{\mathbf{x}^{(d_1)}} e(\gamma(\widetilde{\mathbf{x}};\widetilde{\mathbf{y}}))\right| \ll N(\widehat{\mathbf{x}};\widehat{\mathbf{y}})(P_1\log P_1)^{n_1}.$$

Furthermore, let $M_1(\boldsymbol{\alpha};P_1;P_2;P_1^{-1})$ be the number of integer vectors $\widehat{\mathbf{x}}\in(-P_1,P_1)^{(d_1-1)n_1}$ and $\widetilde{\mathbf{y}}\in(-P_2,P_2)^{d_2n_2}$, such that

$$\|\Gamma(\widehat{\mathbf{x}},\mathbf{e}_l;\widetilde{\mathbf{y}})\| < P_1^{-1}$$

holds for all $1\leq l\leq n_1$. Summing over all $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$ in equation (2.2) gives us the bound

$$|S(\boldsymbol{\alpha})|^{2^{\tilde{d}}} \ll P_1^{n_1(2^{\tilde{d}}-d_1+1)+\varepsilon}P_2^{n_2(2^{\tilde{d}}-d_2)}M_1(\boldsymbol{\alpha};P_1;P_2;P_1^{-1}).$$

The above discussion delivers now the following lemma.

Lemma 2.1. *Let P be a large real number, and $\varepsilon>0$. Then, for some real $\kappa>0$, one has either the upper bound*

$$|S(\boldsymbol{\alpha})| < P_1^{n_1+\varepsilon}P_2^{n_2}P^{-\kappa},$$

or the lower bound

$$M_1(\boldsymbol{\alpha};P_1;P_2;P_1^{-1}) \gg P_1^{n_1(d_1-1)}P_2^{n_2d_2}P^{-2^{\tilde{d}}\kappa}.$$

Next we want to apply the geometry of numbers to $M_1(\boldsymbol{\alpha};P_1;P_2;P_1^{-1})$, similar as done in Birch's work [1] in Lemma 2.3 and Lemma 2.4. For this we need a modified version of a certain lemma from the geometry of numbers which we give in the following section.

3. A lemma from the geometry of numbers

For some integers n_1 and n_2 and real numbers λ_{ij} for $1\leq i\leq n_1$ and $1\leq j\leq n_2$, we consider the linear forms

$$L_i(\mathbf{u})=\sum_{j=1}^{n_2}\lambda_{ij}u_j,$$

and the linear forms corresponding to the transposed matrix of (λ_{ij}) given by

$$L_j^t(\mathbf{u})=\sum_{i=1}^{n_1}\lambda_{ij}u_i.$$

Furthermore, for some real $a>1$ we define $U(Z)$ to be the number of integer tuples $u_1,\ldots,u_{n_2},\ldots,u_{n_1+n_2}$, which satisfy

$$|u_j| < aZ,$$

for $1\leq j\leq n_2$ and

$$|L_i(u_1,\ldots,u_{n_2})-u_{n_2+i}| < a^{-1}Z,$$

for $1\leq i\leq n_1$. Let $U^t(Z)$ be defined analogously with L_i replaced by the linear system L_j^t. Our goal of this section is to establish the following lemma using the geometry of numbers.

Lemma 3.1. *If* $0 < Z_1 \leq Z_2 \leq 1$*, then one has the bound*

$$U(Z_2) \ll \max\left(\left(\frac{Z_2}{Z_1}\right)^{n_2} U(Z_1), \frac{Z_2^{n_2}}{Z_1^{n_1}} a^{n_2-n_1} U^t(Z_1) \right).$$

In the case of $n_1 = n_2$ and symmetric coefficients λ_{ij}, i.e. $\lambda_{ij} = \lambda_{ji}$ for all i, j, this is just Lemma 12.6 from [3]. In our proof we follow mainly the arguments of Davenport in section 12 of [3].

Proof. We start in defining the lattice Γ via the matrix

$$\Lambda = \begin{pmatrix} a^{-1} I_{n_2} & 0 \\ a\lambda & aI_{n_1} \end{pmatrix},$$

where we write I_n for the n-dimensional identity matrix and λ for the $n_1 \times n_2$-matrix with entries λ_{ij}. Let $R_1, \ldots, R_{n_1+n_2}$ be the successive minima of Λ. Furthermore consider the adjoint lattice given by

$$M = (\Lambda^t)^{-1} = \begin{pmatrix} aI_{n_2} & -a\lambda^t \\ 0 & a^{-1} I_{n_1} \end{pmatrix},$$

where λ^t is the transposed matrix of λ. As pointed out by Davenport in section 12 of [3], M has the same successive minima $S_1, \ldots, S_{n_1+n_2}$ as the lattice

$$\widetilde{M} = \begin{pmatrix} a^{-1} I_{n_1} & 0 \\ a\lambda^t & aI_{n_2} \end{pmatrix}.$$

Note that M and Λ are by construction adjoint lattices. Set $b = a^{(n_2-n_1)/(n_1+n_2)}$ and consider the normalised lattices $\Lambda^{\text{nor}} = b\Lambda$ and $M^{\text{nor}} = b^{-1}\widetilde{M}$. Then Λ^{nor} and M^{nor} are adjoint lattices of determinant 1. Let R_i^{nor}, $1 \leq i \leq n_1 + n_2$ and S_i^{nor}, $1 \leq i \leq n_1 + n_2$ be the corresponding succissive minima. Then Mahler's lemma (see for example Lemma 12.5 of [3]) delivers

$$R_k^{\text{nor}} \asymp (S_{n_1+n_2+1-k}^{\text{nor}})^{-1},$$

for all $1 \leq k \leq n_1 + n_2$.

We note that $R_i^{\text{nor}} = bR_i$ and $S_i^{\text{nor}} = b^{-1} S_i$ for all i, and hence we have the relations

$$R_k \asymp S_{n_1+n_2+1-k}^{-1},$$

for all $1 \leq k \leq n_1 + n_2$.

Next let $U_0(Z)$ and $U_0^t(Z)$ be the number of lattice points on Λ and $\widetilde{M}$, whose euclidean norm is bounded by Z. Then one has

$$U_0(Z) \leq U(Z) \leq U_0(\sqrt{n_1+n_2} Z),$$

and the analogous relation holds for U^t and U_0^t. Therefore, we see that it is enough to establish the bound

$$U_0(Z_2) \ll_{n_1,n_2} \max\left(\left(\frac{Z_2}{Z_1}\right)^{n_2} U_0(Z_1), \frac{Z_2^{n_2}}{Z_1^{n_1}} a^{n_2-n_1} U_0^t(Z_1) \right),$$

for all $0 < Z_1 \leq Z_2 \leq \sqrt{n_1 + n_2}$.

For this we first assume that $R_1 \leq Z_1$ and $S_1 \leq Z_1$, and then define the natural numbers μ, ν and ω by

$$R_\nu \leq Z_1 < R_{\nu+1}, \quad R_\mu \leq Z_2 < R_{\mu+1},$$

and

$$S_\omega \leq Z_1 < S_{\omega+1}.$$

Let $U_0^{\mathrm{nor}}(Z)$ be the number of lattice points on Λ^{nor} with euclidean norm bounded by Z. Note that $R_\nu \leq Z_1 < R_{\nu+1}$ is the same as saying that $R_\nu^{\mathrm{nor}} \leq bZ_1 < R_{\nu+1}^{\mathrm{nor}}$, and that one has $U_0(Z) = U_0^{\mathrm{nor}}(bZ)$. Hence Lemma 12.4 of [3] delivers

$$U_0(Z_1) = U_0^{\mathrm{nor}}(bZ_1) \asymp \frac{(bZ_1)^\nu}{R_1^{\mathrm{nor}} \dots R_\nu^{\mathrm{nor}}} = \frac{Z_1^\nu}{R_1 \dots R_\nu}.$$

With the same argument applied to $U_0(Z_2)$ we obtain

$$\frac{U_0(Z_2)}{U_0(Z_1)} \asymp \frac{Z_2^\mu R_1 \dots R_\nu}{Z_1^\nu R_1 \dots R_\mu}.$$

If $\mu \leq n_2$, then we can estimate

$$\frac{U_0(Z_2)}{U_0(Z_1)} \ll \frac{Z_2^\mu}{Z_1^\nu R_{\nu+1} \dots R_\mu} \ll \left(\frac{Z_2}{Z_1}\right)^\mu \ll \left(\frac{Z_2}{Z_1}\right)^{n_2},$$

which is good enough for our lemma. If we have $\mu > n_2$ and $R_{n_2+1} \geq C_1$ for some positive constant C_1 to be chosen later, then we have

$$\frac{Z_2^\mu}{Z_1^\nu R_{\nu+1} \dots R_\mu} \ll \frac{Z_2^{n_2}}{Z_1^{n_2} R_{n_2+1} \dots R_\mu} \ll_{n_1,n_2,C_1} \left(\frac{Z_2}{Z_1}\right)^{n_2},$$

for $\nu \leq n_2$, and

$$\frac{Z_2^\mu}{Z_1^\nu R_{\nu+1} \dots R_\mu} \ll_{C_1} 1 \ll \left(\frac{Z_2}{Z_1}\right)^{n_2},$$

for $\nu > n_2$ using $Z_1 \geq R_{n_2+1} \geq C_1$.

Next assume $\mu > n_2$ and $R_{n_2+1} < C_1$, and note that we have $S_\omega \leq Z_1 \leq \sqrt{n_1+n_2}$. Let c be some positive constant such that $R_{n_2+1} S_{n_1} > c$. Then we obtain $S_{n_1} > \frac{c}{C_1}$. We set $C_1 = c\sqrt{n_1+n_2}^{-1}$, which delivers $S_{n_1} > \sqrt{n_1+n_2}$ and thus $\omega < n_1$. Now consider

$$\frac{U_0(Z_2)}{U_0^t(Z_1)} \asymp \frac{Z_2^\mu S_1 \dots S_\omega}{Z_1^\omega R_1 \dots R_\mu} \asymp \frac{Z_2^\mu}{Z_1^\omega}(S_1 \dots S_\omega)(S_{n_1+n_2+1-\mu} \dots S_{n_1+n_2}). \tag{3.1}$$

We use the relation

$$S_1 \dots S_{n_1+n_2} \asymp b^{n_1+n_2} S_1^{\mathrm{nor}} \dots S_{n_1+n_2}^{\mathrm{nor}} \asymp b^{n_1+n_2}.$$

Hence, if $\omega \leq n_1 + n_2 - \mu$ we can bound the right hand side of equation (3.1) by

$$\ll \frac{Z_2^\mu a^{n_2-n_1}}{Z_1^\omega S_{\omega+1} \dots S_{n_1+n_2-\mu}} \ll \frac{Z_2^{n_2} a^{n_2-n_1}}{Z_1^{n_1+n_2-\mu}} \ll \frac{Z_2^{n_2}}{Z_1^{n_1}} a^{n_2-n_1},$$

since $\mu > n_2$ and $Z_1 \ll 1$. If $\omega > n_1 + n_2 - \mu$, then we obtain in a similar way the bound

$$\begin{aligned}\frac{U_0(Z_2)}{U_0^t(Z_1)} &\ll \frac{Z_2^\mu}{Z_1^\omega} S_{n_1+n_2+1-\mu} \dots S_\omega a^{n_2-n_1} \\ &\ll \frac{Z_2^{n_2}}{Z_1^{n_1}} Z_1^{n_1-\omega} S_{n_1+n_2+1-\mu} \dots S_\omega a^{n_2-n_1} \ll \frac{Z_2^{n_2}}{Z_1^{n_1}} a^{n_2-n_1},\end{aligned}$$

using $S_\omega \leq Z_1 \ll 1$ and $Z_1 \ll 1$.

If $Z_1 < R_1$ or $Z_1 < S_1$ the same computations as above show the inequality which we want to prove, using the observation $U_0(Z_1) = 1$ or $U_0^t(Z_1) = 1$ in these cases. $\square$

4. A form of Weyl's inequality

First we introduce the counting function $M_2(\boldsymbol{\alpha}; P_1; P_2; P^{-1})$ to be the number of integer vectors $\widetilde{\mathbf{x}} \in (-P_1, P_1)^{d_1 n_1}$ and $\widehat{\mathbf{y}} \in (-P_2, P_2)^{(d_2-1)n_2}$ such that

$$\|\Gamma(\widetilde{\mathbf{x}}; \widehat{\mathbf{y}}, \mathbf{e}_l)\| < P^{-1},$$

for $1 \leq l \leq n_2$. Here P is some large real number to be specified later. We need this function for our bounds of $M_1(\boldsymbol{\alpha}; P_1; P_2; P^{-1})$, which we introduced in the last section. We start in writing

$$M_1(\boldsymbol{\alpha}; P_1; P_2; P_1^{-1}) = \sum_{\widehat{\mathbf{x}} \in (-P_1, P_1)^{(d_1-1)n_1}} \sum_{\widehat{\mathbf{y}} \in (-P_2, P_2)^{(d_2-1)n_2}} M_{\widehat{\mathbf{x}}, \widehat{\mathbf{y}}}(P_2, P_1^{-1}),$$

where $M_{\widehat{\mathbf{x}}, \widehat{\mathbf{y}}}(P_2, P_1^{-1})$ is the number of integer vectors $\mathbf{y}^{(d_2)} \in (-P_2, P_2)^{n_2}$ such that

$$\|\Gamma(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{y}^{(d_2)})\| < P_1^{-1},$$

for $1 \leq l \leq n_1$. We apply Lemma 3.1 to the linear forms $\Gamma(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{y}^{(d_2)})$ in the variables $\mathbf{y}^{(d_2)}$. Let $0 < \theta_2 \leq 1$ be fixed. We choose the parameters Z_1, Z_2 and a such that

$$P_2 = aZ_2 \qquad P_2^{\theta_2} = aZ_1$$
$$P_1^{-1} = a^{-1} Z_2.$$

This gives $a^{-1} Z_1 = P_1^{-1} P_2^{-1+\theta_2}$. Furthermore note that $Z_2 \leq 1$ since we have $P_2 \leq P_1$.

Recall that Lemma 3.1 gives a bound of the form

$$U(Z_2) \ll \max \left(\left(\frac{aZ_2}{aZ_1} \right)^{n_2} U(Z_1), \frac{(aZ_2)^{n_2}}{(aZ_1)^{n_1}} U^t(Z_1) \right).$$

Hence, we have

$$\begin{aligned}M_{\widehat{\mathbf{x}}, \widehat{\mathbf{y}}}(P_2, P_1^{-1}) \ll \max(&P_2^{n_2(1-\theta_2)} M_{\widehat{\mathbf{x}}, \widehat{\mathbf{y}}}(P_2^{\theta_2}, P_1^{-1} P_2^{-1+\theta_2}), \\ &P_2^{n_2 - n_1\theta_2} M_{\widehat{\mathbf{x}}, \widehat{\mathbf{y}}}^t(P_2^{\theta_2}, P_1^{-1} P_2^{-1+\theta_2})),\end{aligned}$$

where $M^t_{\widehat{\mathbf{x}},\widehat{\mathbf{y}}}$ counts the solutions of the corresponding transposed linear system as in section 5. For this we write

$$\Gamma(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{y}^{(d_2)}) = \sum_{m=1}^{n_2} \lambda_{lm} y_m^{(d_2)},$$

with

$$\lambda_{lm} = \Gamma(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{e}_m).$$

Still with the notation from section 5 we have

$$L_m^t(\mathbf{y}^{(d_2)}) = \sum_{l=1}^{n_2} \lambda_{lm} y_l^{(d_2)} = \Gamma(\widehat{\mathbf{x}}, \mathbf{y}^{(d_2)}; \widehat{\mathbf{y}}, \mathbf{e}_m).$$

Therefore, we see that $M^t_{\widehat{\mathbf{x}},\widehat{\mathbf{y}}}(P_2^{\theta_2}, P_1^{-1}P_2^{-1+\theta_2})$ counts the number of integer vectors $\mathbf{z} \in (-P_2^{\theta_2}, P_2^{\theta_2})^{n_1}$ with

$$\|\Gamma(\widehat{\mathbf{x}}, \mathbf{z}; \widehat{\mathbf{y}}, \mathbf{e}_m)\| < P_1^{-1}P_2^{-1+\theta_2},$$

for $1 \leq m \leq n_2$. Taking the sum over all the contributions of admissible $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$ we obtain

$$M_1(\boldsymbol{\alpha}; P_1; P_2; P_1^{-1}) \ll S_1 P_2^{n_2(1-\theta_2)} + S_2 P_2^{n_2 - n_1\theta_2}.$$

Here S_1 counts all integer vectors $\widehat{\mathbf{x}} \in (-P_1, P_1)^{(d_1-1)n_1}$ and $\widehat{\mathbf{y}} \in (-P_2, P_2)^{(d_2-1)n_2}$ and $\mathbf{z} \in (-P_2^{\theta_2}, P_2^{\theta_2})^{n_2}$ with

$$\|\Gamma(\widehat{\mathbf{x}}, \mathbf{e}_l; \widehat{\mathbf{y}}, \mathbf{z})\| < P_1^{-1}P_2^{-1+\theta_2},$$

for $1 \leq l \leq n_1$, and S_2 is the number of $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$ in the same region and $\mathbf{z} \in (-P_2^{\theta_2}, P_2^{\theta_2})^{n_1}$ such that

$$\|\Gamma(\widehat{\mathbf{x}}, \mathbf{z}; \widehat{\mathbf{y}}, \mathbf{e}_l)\| < P_1^{-1}P_2^{-1+\theta_2},$$

for $1 \leq l \leq n_2$.

Next we define θ_1 by the relation $P_1^{\theta_1} = P_2^{\theta_2}$ and note that we have $0 < \theta_1 \leq 1$ by the assumption on P_1 and P_2. For convenience we write $P_1^{\theta_1} = P^{\theta}$ for some real number θ and some $P \geq 2$. Now we iterate the above procedure with repect to all the vectors from $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$. This delivers the bound

$$\begin{aligned} M_1(\boldsymbol{\alpha}; &P_1; P_2; P_1^{-1}) \ll P_1^{n_1(d_1-1)} P_2^{n_2 d_2} P^{-\theta(n_1 d_1 + n_2 d_2)} \\ &\times (P^{n_1\theta} M_1(\boldsymbol{\alpha}; P^\theta; P^\theta; P_1^{-d_1}P_2^{-d_2}P^{\theta(\tilde{d}+1)}) + P^{n_2\theta} M_2(\boldsymbol{\alpha}; P^\theta; P^\theta; P_1^{-d_1}P_2^{-d_2}P^{\theta(\tilde{d}+1)})) \end{aligned}$$

In combination with Lemma 2.1 we obtain the following result.

Lemma 4.1. *Under the above assumptions one has either the upper bound*

$$|S(\boldsymbol{\alpha})| < P_1^{n_1+\varepsilon} P_2^{n_2} P^{-\kappa},$$

or the lower bound

$$M_i(\boldsymbol{\alpha}; P^\theta; P^\theta; P_1^{-d_1}P_2^{-d_2}P^{\theta(\tilde{d}+1)}) \gg P^{\theta(n_1 d_1 + n_2 d_2) - \theta n_i} P^{-2^{\tilde{d}}\kappa},$$

for $i = 1$ or $i = 2$.

Next we proceed similarly as in Birch's work [1]. We write

$$\Gamma(\widetilde{\mathbf{x}};\widetilde{\mathbf{y}}) = \sum_{i=1}^{R} \alpha_i \Gamma_i(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}}),$$

with

$$\Gamma_i(\widetilde{\mathbf{x}};\widetilde{\mathbf{y}}) = d_1! d_2! \sum_{\mathbf{j}} \sum_{\mathbf{k}} F^{(i)}_{\mathbf{j},\mathbf{k}} x^{(1)}_{j_1} \dots x^{(d_1)}_{j_{d_1}} y^{(1)}_{k_1} \dots y^{(d_2)}_{k_{d_2}}.$$

Suppose that we have some integer vectors $\widehat{\mathbf{x}} \in (-P^\theta, P^\theta)^{n_1(d_1-1)}$ and $\widetilde{\mathbf{y}} \in (-P^\theta, P^\theta)^{n_2 d_2}$ counted by $M_1(\boldsymbol{\alpha}; P^\theta; P^\theta; P_1^{-d_1} P_2^{-d_2} P^{\theta(\tilde{d}+1)})$ such that the matrix

$$(\Gamma_i(\widehat{\mathbf{x}},\mathbf{e}_l;\widetilde{\mathbf{y}}))_{\substack{1\leq i\leq R\\ 1\leq l\leq n_1}}$$

has full rank. Without loss of generality we may assume that the leading $R\times R$ minor has full rank. Our next goal is to show that in this case the α_i are well approximated by rational numbers. For this we write

$$\Gamma(\widehat{\mathbf{x}},\mathbf{e}_l;\widetilde{\mathbf{y}}) = \tilde{a}_l + \widetilde{\delta}_l,$$

for $1\leq l\leq n_1$, with some integers $\tilde{a}_l$ and real $\widetilde{\delta}_l$ with $|\widetilde{\delta}_l| < P_1^{-d_1} P_2^{-d_2} P^{\theta(\tilde{d}+1)}$. Next let q be the absolute value of the determinant of the matrix $(\Gamma_i(\widehat{\mathbf{x}},\mathbf{e}_l;\widetilde{\mathbf{y}}))_{1\leq i,l\leq R}$, and note that we have

$$q \ll P^{R\theta(\tilde{d}+1)}.$$

Using the formula for the adjoint matrix of our matrix under consideration we obtain

$$\alpha_i = q^{-1}(a_i + \delta_i),$$

for $1\leq i\leq R$ with some integers a_i and with

$$|\delta_i| \ll P^{(R-1)\theta(\tilde{d}+1)} \max_l |\widetilde{\delta}_l|.$$

Thus, we obtain the approximation

$$|q\alpha_i - a_i| \ll P_1^{-d_1} P_2^{-d_2} P^{R\theta(\tilde{d}+1)},$$

for $1\leq i\leq R$.

We have now established the following lemma.

Lemma 4.2. *There is some positive constant C such that the following holds. Let $P_2 \leq P_1$ and P some real number larger than 2. Let $0<\theta_2\leq 1$ and write $P_2^{\theta_2} = P^\theta$. Then at least one of the following alternatives hold.*

i) One has the upper bound $|S(\boldsymbol{\alpha})| < P_1^{n_1+\varepsilon} P_2^{n_2} P^{-\kappa}$.

ii) There exist integers $1\leq q\leq P^{R(\tilde{d}+1)\theta}$ and $a_1,\dots,a_R$ with

$$\gcd(q,a_1,\dots,a_R) = 1,$$

and

$$2|q\alpha_i - a_i| \leq P_1^{-d_1} P_2^{-d_2} P^{R(\tilde{d}+1)\theta},$$

for $1\leq i\leq R$.

iii) The number of integer vectors $\widehat{\mathbf{x}} \in (-P^\theta, P^\theta)^{n_1(d_1-1)}$ *and* $\widetilde{\mathbf{y}} \in (-P^\theta, P^\theta)^{n_2 d_2}$ *such that*

$$\operatorname{rank}(\Gamma_i(\widehat{\mathbf{x}}, \mathbf{e}_l; \widetilde{\mathbf{y}})) < R \tag{4.1}$$

is bounded below by

$$\geq C(P^\theta)^{n_1(d_1-1)+n_2 d_2 - 2^{\tilde{d}}\kappa/\theta}.$$

iv) The number of integer vectors $\widetilde{\mathbf{x}} \in (-P^\theta, P^\theta)^{n_1 d_1}$ *and* $\widehat{\mathbf{y}} \in (-P^\theta, P^\theta)^{n_2(d_2-1)}$ *such that*

$$\operatorname{rank}(\Gamma_i(\widetilde{\mathbf{x}}; \widehat{\mathbf{y}}, \mathbf{e}_l)) < R \tag{4.2}$$

is bounded below by

$$\geq C(P^\theta)^{n_1 d_1 + n_2(d_2-1) - 2^{\tilde{d}}\kappa/\theta}.$$

We note that the constant C is independent of θ_2.

Assume that alternative iii) of the above lemma holds. Let $\mathcal{L}_1$ be the affine variety defined by equation (4.1) in affine $n_1(d_1-1)+n_2 d_2$-space. As in Birch's work [1], section 3, the condition iii) implies the lower bound

$$\dim \mathcal{L}_1 \geq n_1(d_1-1) + n_2 d_2 - 2^{\tilde{d}}\kappa/\theta.$$

Recall that the affine variety V_1^* (see equation (1.2) in $\mathbb{A}_{\mathbb{C}}^{n_1+n_2}$ is given by

$$\operatorname{rank}\left(\frac{\partial F_i}{\partial x_j}\right)_{\substack{1\leq i\leq R\\ 1\leq j\leq n_1}} < R.$$

Furthermore, let $\mathcal{D}$ be the linear subspace given by

$$\mathbf{x}^{(1)} = \ldots = \mathbf{x}^{(d_1-1)} \text{ and } \mathbf{y}^{(1)} = \ldots = \mathbf{y}^{(d_2)},$$

in affine $n_1(d_1-1)+n_2 d_2$-space. Considering these as varieties over the algebraically closed field $\mathbb{C}$ one has

$$\dim \mathcal{L}_1 \cap \mathcal{D} \geq \dim \mathcal{L}_1 - n_2(d_2-1) - n_1(d_1-2).$$

Since $\mathcal{L}_1 \cap \mathcal{D}$ projects onto V_1^*, condition iii) above implies

$$\dim V_1^* \geq n_1 + n_2 - 2^{\tilde{d}}\kappa/\theta.$$

Similarly, we note that condition iv) of Lemma 4.2 implies

$$\dim V_2^* \geq n_1 + n_2 - 2^{\tilde{d}}\kappa/\theta.$$

Define K by

$$2^{\tilde{d}}K = \min\{n_1+n_2-\dim V_1^*, n_1+n_2-\dim V_2^*\}.$$

Furthermore we set $P = P_1^{d_1}P_2^{d_2}$ for the rest of this paper. Note that this gives the relations

$$\theta = (bd_1+d_2)^{-1}\theta_2,$$

and

$$\theta_1 = b^{-1}\theta_2.$$

Next we define $\mathfrak{M}(\theta)$ to be the set of $\boldsymbol{\alpha} \in [0,1]^R$ such that $\boldsymbol{\alpha}$ satisfies condition ii) of Lemma 4.2. With this notation we can state our final lemma of this section, which is a direct consequence of Lemma 4.2.

Lemma 4.3. *Let $0 < \theta \leq (bd_1 + d_2)^{-1}$ and assume $\varepsilon > 0$. Then one has for some real vector $\boldsymbol{\alpha} \in \mathbb{R}^R$ either $\boldsymbol{\alpha} \in \mathfrak{M}(\theta)$ modulo 1 or the upper bound*

$$|S(\boldsymbol{\alpha})| \ll P_1^{n_1} P_2^{n_2} P^{-K\theta+\varepsilon}.$$

5. Circle method

In this section we set up the circle method to get an asymptotic formula for $N(P_1, P_2)$ mainly following Birch's work [1]. We note that by orthogonality we have

$$N(P_1, P_2) = \int_{[0,1]^R} S(\boldsymbol{\alpha}) \, \mathrm{d}\boldsymbol{\alpha}. \tag{5.1}$$

In the following we assume that we have

$$K > \max\{R(R+1)(\tilde{d}+1), R(bd_1 + d_2)\}. \tag{5.2}$$

Next we choose positive and real δ and ϑ_0 in such a way that the following conditions are satisfied

$$K - R(R+1)(\tilde{d}+1) > 2\delta\vartheta_0^{-1}, \tag{5.3}$$

$$K > (2\delta + R)(bd_1 + d_2), \tag{5.4}$$

and

$$1 > (bd_1 + d_2)R(\tilde{d}+1)\vartheta_0(2R+3) + \delta(bd_1 + d_2). \tag{5.5}$$

Note that the parameters δ and ϑ_0 may depend on b. Now we use the results of the last section to show that the contribution of those $\boldsymbol{\alpha}$ which are not in $\mathfrak{M}(\vartheta_0)$ is neglegible in equation (5.1). This is done in the following lemma.

Lemma 5.1. *One has*

$$\int_{\boldsymbol{\alpha} \notin \mathfrak{M}(\vartheta_0)} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha} = O(P_1^{n_1} P_2^{n_2} P^{-R-\delta}).$$

Proof. We choose a sequence of ϑ_i with

$$\vartheta_T > \vartheta_{T-1} > \ldots > \vartheta_1 > \vartheta_0 > 0,$$

and

$$\vartheta_T \leq (bd_1 + d_2)^{-1} \quad \text{and} \quad \vartheta_T K > 2\delta + R.$$

Note that this is possible by equation (5.4). Furthermore we choose our ϑ_i in such a way that they satisfy

$$\frac{1}{2}\delta > R(R+1)(\tilde{d}+1)(\vartheta_{t+1} - \vartheta_t),$$

for $0 \leq t < T$. We certainly can achieve this with $T \ll P^{\delta/2}$.

Now we consider the contribution of those $\boldsymbol{\alpha}$, which do not belong to $\mathfrak{M}(\vartheta_T)$. By Lemma 4.3 we have

$$\begin{aligned} \int_{\boldsymbol{\alpha} \notin \mathfrak{M}(\vartheta_T)} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha} &\ll P_1^{n_1} P_2^{n_2} P^{-K\vartheta_T+\varepsilon} \\ &\ll P_1^{n_1} P_2^{n_2} P^{-R-\delta}. \end{aligned}$$

For some $\theta > 0$ we can estimate the measure of $\mathfrak{M}(\theta)$ by

$$\begin{aligned}\operatorname{meas}(\mathfrak{M}(\theta)) &\ll \sum_{q \leq P^{R(\tilde{d}+1)\theta}} \sum_{\mathbf{a}} q^{-R} P_1^{-d_1 R} P_2^{-d_2 R} P^{R^2(\tilde{d}+1)\theta} \\ &\ll P^{-R+R(R+1)(\tilde{d}+1)\theta}.\end{aligned}$$

This estimate together with Lemma 4.3 delivers the bound

$$\int_{\boldsymbol{\alpha} \in \mathfrak{M}(\vartheta_{t+1}) \setminus \mathfrak{M}(\vartheta_t)} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha} \ll P_1^{n_1} P_2^{n_2} P^{-K\vartheta_t + \varepsilon - R + R(R+1)(\tilde{d}+1)\vartheta_{t+1}}.$$

Since we have the inequality

$$-K\vartheta_t + R(R+1)(\tilde{d}+1)\vartheta_{t+1} \leq \frac{1}{2}\delta + \vartheta_t(-K + R(R+1)(\tilde{d}+1)) \leq \frac{1}{2}\delta - 2\delta,$$

we finally obtain the estimate

$$\int_{\boldsymbol{\alpha} \in \mathfrak{M}(\vartheta_{t+1}) \setminus \mathfrak{M}(\vartheta_t)} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha} \ll P_1^{n_1} P_2^{n_2} P^{-R-3\delta/2},$$

for $0 \leq t < T$, which is enough to prove the lemma. □

Next we turn towards the contribution of the major arcs. In order to obtain nicer formulas, we first define some modified major arcs. For some q and $0 \leq a_i < q$ let $\mathfrak{M}'_{\mathbf{a},q}(\theta)$ be the set of $\boldsymbol{\alpha} \in [0,1]^R$ such that

$$|q\alpha_i - a_i| \leq q P^{-1+R(\tilde{d}+1)\theta},$$

for $1 \leq i \leq R$. In the same way as before we set

$$\mathfrak{M}'(\theta) = \bigcup_{1 \leq q \leq P^{R(\tilde{d}+1)\theta}} \bigcup_{\mathbf{a}} \mathfrak{M}'_{\mathbf{a},q}(\theta),$$

where the union for the $\mathbf{a}$ is over all $0 \leq a_i < q$ with $\gcd(q, a_1, \ldots, a_R) = 1$. We note that the $\mathfrak{M}'_{\mathbf{a},q}(\theta)$ are disjoint if θ is sufficiently small. If we have in the above union some

$$\boldsymbol{\alpha} \in \mathfrak{M}'_{\mathbf{a},q}(\theta) \cap \mathfrak{M}'_{\tilde{\mathbf{a}},\tilde{q}}(\theta),$$

for distinct $\mathbf{a}, q$ and $\tilde{\mathbf{a}}, \tilde{q}$, then there is some $1 \leq i \leq R$ such that

$$\frac{1}{q\tilde{q}} \leq \left| \frac{a_i}{q} - \frac{\tilde{a}_i}{\tilde{q}} \right| \leq 2P^{-1+R(\tilde{d}+1)\theta}.$$

This is impossible for large P and $\theta < 1/(3R(\tilde{d}+1))$. By equation (5.5) we see that our major arcs $\mathfrak{M}'(\vartheta_0)$ are disjoint. Thus, we have the following lemma, which is a direct consequence of Lemma 5.1 and equation (5.1).

Lemma 5.2. *One has*

$$N(P_1, P_2) = \sum_{1 \leq q \leq P^{R(\tilde{d}+1)\vartheta_0}} \sum_{\mathbf{a}} \int_{\mathfrak{M}'_{\mathbf{a},q}(\vartheta_0)} S(\boldsymbol{\alpha}) \, \mathrm{d}\boldsymbol{\alpha} + O(P_1^{n_1} P_2^{n_2} P^{-R-\delta}),$$

where the second sum is over all $0 \leq a_i < q$ for $1 \leq i \leq R$, such that

$$\gcd(q, a_1, \ldots, a_R) = 1.$$

Our next goal is to obtain an approximation for $S(\boldsymbol{\alpha})$ on the major arcs. For convenience we write in the following $\eta = R(\tilde{d}+1)\vartheta_0$. Furthermore, for some $\boldsymbol{\alpha} \in \mathfrak{M}'_{\mathbf{a},q}(\vartheta_0)$ we write $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta}$ with

$$|\beta_i| \leq P^{-1+\eta},$$

for $1 \leq i \leq R$. We introduce the notation

$$S_{\mathbf{a},q} = \sum_{\mathbf{x},\mathbf{y}} e\left(\sum_{i=1}^{R} a_i F_i(\mathbf{x},\mathbf{y})/q\right),$$

where $\mathbf{x}$ and $\mathbf{y}$ run through a complete set of residues modulo q. Let

$$I(\mathbf{u}) = \int_{\mathcal{B}_1\times\mathcal{B}_2} e\left(\sum_{i=1}^{R} u_i F_i(\mathbf{v};\mathbf{w})\right) \mathrm{d}\mathbf{v}\,\mathrm{d}\mathbf{w},$$

for some real vector $\mathbf{u} = (u_1,\ldots,u_R)$. Now we have introduced all the notation we need to state our next lemma.

Lemma 5.3. *Let $\boldsymbol{\alpha} \in \mathfrak{M}'_{\mathbf{a},q}(\vartheta_0)$ and $q \leq P^{\eta}$. Then one has*

$$S(\boldsymbol{\alpha}) = P_1^{n_1}P_2^{n_2}q^{-n_1-n_2}S_{\mathbf{a},q}I(P\boldsymbol{\beta}) + O(P_1^{n_1}P_2^{n_2}P^{2\eta}P_2^{-1}).$$

Proof. In the sum $S(\boldsymbol{\alpha})$ we write $\mathbf{x} = \mathbf{z}^{(1)} + q\mathbf{x}'$ and $\mathbf{y} = \mathbf{z}^{(2)} + q\mathbf{y}'$, with $0 \leq z_i^{(1)} < q$ and $0 \leq z_i^{(2)} < q$ for all $1 \leq i \leq n$. Then we obtain

$$\begin{aligned} S(\boldsymbol{\alpha}) &= \sum_{\mathbf{x}\in P_1\mathcal{B}_1}\sum_{\mathbf{y}\in P_2\mathcal{B}_2} e\left(\sum_{i=1}^{R}\alpha_i F_i(\mathbf{x};\mathbf{y})\right) \\ &= \sum_{\mathbf{z}^{(1)}}\sum_{\mathbf{z}^{(2)}} e\left(\sum_{i=1}^{R} a_i F_i(\mathbf{z}^{(1)};\mathbf{z}^{(2)})/q\right) S_3(\mathbf{z}^{(1)},\mathbf{z}^{(2)}), \end{aligned}$$

with the sum

$$S_3(\mathbf{z}^{(1)},\mathbf{z}^{(2)}) = \sum_{\mathbf{x}'}\sum_{\mathbf{y}'} e\left(\sum_{i=1}^{R}\beta_i F_i(q\mathbf{x}'+\mathbf{z}^{(1)};q\mathbf{y}'+\mathbf{z}^{(2)})\right),$$

where the integer vectors $\mathbf{x}'$ run through a range such that $q\mathbf{x}'+\mathbf{z}^{(1)} \in P_1\mathcal{B}_1$ and for $\mathbf{y}'$ analogously.

Consider some vectors $\mathbf{x}',\mathbf{x}''$ and $\mathbf{y}',\mathbf{y}''$ with

$$\max_{1\leq i\leq n_1}|x_i' - x_i''| \leq 2,$$

and

$$\max_{1\leq i\leq n_2}|y_i' - y_i''| \leq 2.$$

In this case one has

$$\begin{aligned} |F_i(q\mathbf{x}'+\mathbf{z}^{(1)};q\mathbf{y}'+\mathbf{z}^{(2)}) - F_i(q\mathbf{x}''+\mathbf{z}^{(1)};q\mathbf{y}''+\mathbf{z}^{(2)})| &\ll qP_1^{d_1-1}P_2^{d_2} + qP_1^{d_1}P_2^{d_2-1} \\ &\ll qP_1^{d_1}P_2^{d_2-1}. \end{aligned}$$

We replace the sum in S_3 with an integral and obtain

$$S_3 = \int_{q\tilde{\mathbf{v}}\in P_1\mathcal{B}_1}\int_{q\tilde{\mathbf{w}}\in P_2\mathcal{B}_2} e\left(\sum_{i=1}^R \beta_i F_i(q\tilde{\mathbf{v}};q\tilde{\mathbf{w}})\right) \mathrm{d}\tilde{\mathbf{v}}\,\mathrm{d}\tilde{\mathbf{w}}$$
$$+ O\left(\sum_{i=1}^R |\beta_i| q P_1^{d_1}P_2^{d_2-1}\left(\frac{P_1}{q}\right)^{n_1}\left(\frac{P_2}{q}\right)^{n_2} + \left(\frac{P_1}{q}\right)^{n_1}\left(\frac{P_2}{q}\right)^{n_2-1}\right).$$

A variable substitution $\mathbf{v} = qP_1^{-1}\widetilde{\mathbf{v}}$ and $\mathbf{w} = qP_2^{-1}\widetilde{\mathbf{w}}$ in the integral leads to

$$\begin{aligned} S_3 &= P_1^{n_1}P_2^{n_2}q^{-(n_1+n_2)}\int_{\mathbf{v}\in\mathcal{B}_1}\int_{\mathbf{w}\in\mathcal{B}_2} e\left(\sum_{i=1}^R P_1^{d_1}P_2^{d_2}\beta_i F_i(\mathbf{v};\mathbf{w})\right)\mathrm{d}\mathbf{v}\,\mathrm{d}\mathbf{w} \\ &+ O(q^{-n_1-n_2+1}P^\eta P_2^{-1}P_1^{n_1}P_2^{n_2} + q^{-n_1-n_2+1}P_1^{n_1}P_2^{n_2-1}) \\ &= P_1^{n_1}P_2^{n_2}q^{-n_1-n_2}I(P\boldsymbol{\beta}) + O(P_1^{n_1}P_2^{n_2}P^\eta P_2^{-1}q^{-n_1-n_2+1}). \end{aligned}$$

Summing over $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$ we finally obtain the approximation

$$S(\boldsymbol{\alpha}) = P_1^{n_1}P_2^{n_2}q^{-n_1-n_2}S_{\mathbf{a},q}I(P\boldsymbol{\beta}) + O(P_1^{n_1}P_2^{n_2}P^{2\eta}P_2^{-1}),$$

as desired. □

Now we use the approximation of Lemma 5.3 to evaluate the sum over the major arcs from Lemma 5.2. This leads to

$$\begin{aligned} N(P_1,P_2) =& P_1^{n_1}P_2^{n_2}\sum_{1\leq q\leq P^\eta} q^{-n_1-n_2}\sum_{\mathbf{a}} S_{\mathbf{a},q}\int_{|\boldsymbol{\beta}|\leq P^{-1+\eta}} I(P\boldsymbol{\beta})\,\mathrm{d}\boldsymbol{\beta} \\ &+ O(P_1^{n_1}P_2^{n_2}P^{2\eta}P_2^{-1}\mathrm{meas}(\mathfrak{M}'(\vartheta_0))). \end{aligned}$$

The measure of these major arcs is bounded by

$$\mathrm{meas}(\mathfrak{M}'(\vartheta_0)) \ll \sum_{q\leq P^\eta} q^R P^{-R+\eta R} \ll P^{-R+\eta(2R+1)}.$$

We define the sum

$$\mathfrak{S}(P^\eta) = \sum_{1\leq q\leq P^\eta} q^{-n_1-n_2}\sum_{\mathbf{a}} S_{\mathbf{a},q},$$

where the second sum is as before over all tuples $0\leq a_i<q$ with $\gcd(q,a_1,\dots,a_R) = 1$, and we define the integral

$$J(P^\eta) = \int_{|\boldsymbol{\beta}|\leq P^\eta} I(\boldsymbol{\beta})\,\mathrm{d}\boldsymbol{\beta}.$$

With this notation we obtain

$$\begin{aligned} N(P_1,P_2) &= P_1^{n_1}P_2^{n_2}P^{-R}\mathfrak{S}(P^\eta)\int_{|\boldsymbol{\beta}|\leq P^\eta} I(\boldsymbol{\beta})\,\mathrm{d}\boldsymbol{\beta} + O(P_1^{n_1}P_2^{n_2}P^{-R}P_2^{-1}P^{\eta(2R+3)}) \\ &= P_1^{n_1}P_2^{n_2}P^{-R}\mathfrak{S}(P^\eta)J(P^\eta) + O(P_1^{n_1}P_2^{n_2}P^{-R+\eta(2R+3)-1/(bd_1+d_2)}). \end{aligned}$$

The error term is bounded by $O(P_1^nP_2^nP^{-R-\delta})$ if we have

$$\tfrac{1}{bd_1+d_2} > \eta(2R+3)+\delta,$$

which is just equation (5.5). Thus, we have obtained the following asymptotic for $N(P_1, P_2)$.

Lemma 5.4. *Assume that equation (5.2) holds and let δ and ϑ_0 be chosen as at the beginning of this section. Then one has*

$$N(P_1, P_2) = P_1^{n_1} P_2^{n_2} P^{-R} \mathfrak{S}(P^\eta) J(P^\eta) + O(P_1^{n_1} P_2^{n_2} P^{-R-\delta}).$$

Next we consider the terms $\mathfrak{S}(P^\eta)$ and $J(P^\eta)$ separately. First we define the singular series,

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\mathbf{a}} q^{-(n_1+n_2)} S_{\mathbf{a},q}, \tag{5.6}$$

if this series exists. The following lemma shows that this is the case, and that $\mathfrak{S}$ is absolutely convergent.

Lemma 5.5. *The series $\mathfrak{S}$ is absolutely convergent and one has*

$$|\mathfrak{S}(Q) - \mathfrak{S}| \ll Q^{-\delta/\eta},$$

for any large real number Q.

Proof. First we need an estimate for the sums $S_{\mathbf{a},q}$. For this we note that we have

$$S_{\mathbf{a},q} = S(\boldsymbol{\alpha}),$$

if we set $\mathcal{B}_1 = [0,1)^{n_1}$, $\mathcal{B}_2 = [0,1)^{n_2}$ and $P_1 = P_2 = q$ and $\boldsymbol{\alpha} = \mathbf{a}/q$. We define θ by

$$(d_1 + d_2) R (\tilde{d} + 1)\theta = 1 - \varepsilon,$$

for some $\varepsilon > 0$. Then we claim that $\mathbf{a}/q$ cannot lie inside the major arcs $\mathfrak{M}(\theta)$, if we assume $\gcd(q, a_1, \ldots, a_R) = 1$. Otherwise we would have some integers q' and $\mathbf{a}'$ with

$$1 \leq q' \leq q^{(d_1+d_2)R(\tilde{d}+1)\theta},$$

and

$$2|q'a_i - a_i'q| \leq q q^{-d_1} q^{-d_2} q^{(d_1+d_2)R(\tilde{d}+1)\theta},$$

for all $1 \leq i \leq R$, which is impossible. Therefore Lemma 4.3 delivers

$$\begin{aligned} |S_{\mathbf{a},q}| &\ll q^{n_1+n_2} q^{-K(d_1+d_2)[(d_1+d_2)R(\tilde{d}+1)]^{-1}+\varepsilon} \\ &\ll q^{n_1+n_2-K/(R(\tilde{d}+1))+\varepsilon}. \end{aligned}$$

With equation (5.3) this leads to the bound

$$|S_{\mathbf{a},q}| \ll q^{n_1+n_2-R-1-\delta/\eta}.$$

Now we can estimate the desired series

$$\sum_{q>Q} \sum_{\mathbf{a}} q^{-n_1-n_2} |S_{\mathbf{a},q}| \ll \sum_{q>Q} q^{-1-\delta/\eta} \ll Q^{-\delta/\eta},$$

which proves both claims of the lemma. □

Similarly as for the singular series, we define the singular integral

$$J = \int_{\boldsymbol{\beta} \in \mathbb{R}^R} I(\boldsymbol{\beta})\, \mathrm{d}\boldsymbol{\beta}, \tag{5.7}$$

if this exists.

Lemma 5.6. *The singular integral J is absolutely convergent and we have*

$$|J - J(\Phi)| \ll \Phi^{-1},$$

for any large positive real number Φ.

Proof. For convenience of notation we set $B = \max_i |\beta_i|$ for some real vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_R)$, and assume $B \geq 2$. Set $\theta = \vartheta_0$ as we have chosen it at the beginning of this section and define P by

$$2B = P^{R(\tilde{d}+1)\theta}.$$

Then we have $P^{-1}\boldsymbol{\beta} \in \mathfrak{M}_{0,1}(\theta)$, since

$$2|P^{-1}\beta_i| \leq P^{-1}P^{R(\tilde{d}+1)\theta},$$

for all $1 \leq i \leq R$. Then Lemma 5.3 delivers

$$S(P^{-1}\boldsymbol{\beta}) = P_1^{n_1}P_2^{n_2}I(\boldsymbol{\beta}) + O(P_1^{n_1}P_2^{n_2}P^{2R(\tilde{d}+1)\theta}P_2^{-1}). \tag{5.8}$$

Furthermore $P^{-1}\boldsymbol{\beta}$ lies by construction on the boundary of $\mathfrak{M}(\theta)$, which are disjoint by Lemma 4.1 of Birch's paper [1]. Thus, our Lemma 4.3 gives the bound

$$|S(P^{-1}\boldsymbol{\beta})| \ll P_1^{n_1}P_2^{n_2}P^{-K\theta+\varepsilon}.$$

Together with equation (5.8) this implies

$$|I(\boldsymbol{\beta})| \ll P^{-K\vartheta_0+\varepsilon} + P^{2R(\tilde{d}+1)\theta-1/(bd_1+d_2)}.$$

From equation (5.5) we see that

$$\tfrac{1}{bd_1+d_2} - 2R(\tilde{d}+1)\vartheta_0 > 2R(R+1)(\tilde{d}+1)\vartheta_0 + \delta,$$

which implies

$$P^{2R(\tilde{d}+1)\theta-1/(bd_1+d_2)} \ll B^{-2R}.$$

In the same way we see that equation (5.3) gives

$$P^{-K\vartheta_0+\varepsilon} \ll B^{-R-1},$$

such that we have

$$|I(\boldsymbol{\beta})| \ll (\max_i |\beta_i|)^{-R-1}.$$

Now we can use this bound to estimate the integral

$$\int_{\Phi_1 \leq B \leq \Phi_2} |I(\boldsymbol{\beta})|\, \mathrm{d}\boldsymbol{\beta} \ll \int_{\Phi_1 \leq B \leq \Phi_2} B^{R-1}B^{-R-1}\, \mathrm{d}B \ll \Phi_1^{-1}.$$

This shows that J is absolutely convergent and also that the second assertion of the lemma holds. $\square$

6. Conclusions

Before we finish our proof of Theorem 1.1, we give an alternative representation of the singular integral, following Schmidt's work [6]. For this we define the function

$$\psi(z) = \begin{cases} 1 - |z| & \text{for } |z| \leq 1, \\ 0 & \text{for } |z| > 1, \end{cases}$$

and for $T > 0$ we set $\psi_T(z) = T\psi(Tz)$. Furthermore, for some vector $\mathbf{z} = (z_1, \ldots, z_R)$ we define

$$\psi_T(\mathbf{z}) = \psi_T(z_1) \cdot \ldots \cdot \psi_T(z_R).$$

With this notation we define

$$\tilde{J}_T = \int_{\mathcal{B}_1 \times \mathcal{B}_2} \psi_T(\mathbf{F}(\boldsymbol{\xi}^{(1)}; \boldsymbol{\xi}^{(2)})) \, \mathrm{d}\boldsymbol{\xi}^{(1)} \, \mathrm{d}\boldsymbol{\xi}^{(2)},$$

and

$$\tilde{J} = \lim_{T \to \infty} \tilde{J}_T,$$

if the limit exists.

Proof of Theorem 1.1. Note that the assumptions of Theorem 1.1 imply that equation (5.2) holds. Hence, by Lemma 5.4 we have

$$N(P_1, P_2) = P_1^{n_1} P_2^{n_2} P^{-R} \mathfrak{S}(P^\eta) J(P^\eta) + O(P_1^{n_1} P_2^{n_2} P^{-R-\delta}).$$

Together with Lemma 5.5 and Lemma 5.6 this gives

$$N(P_1, P_2) = P_1^{n_1} P_2^{n_2} P^{-R} \mathfrak{S} J + O(P_1^{n_1} P_2^{n_2} P^{-R-\delta}),$$

which already proves the first part of the theorem.

As usual, the singular series $\mathfrak{S}$ factorizes as $\mathfrak{S} = \prod_p \mathfrak{S}_p$, where the product is over all primes p, and

$$\mathfrak{S}_p = \sum_{l=1}^{\infty} \sum_{\mathbf{a}} p^{-(n_1+n_2)l} S_{\mathbf{a}, p^l},$$

where the sum over $\mathbf{a}$ is over all $0 \leq a_i < p^l$ with $\gcd(a_1, \ldots, a_R, p) = 1$. We know in a relatively general context that $\mathfrak{S} > 0$ if the $F_i(\mathbf{x}; \mathbf{y})$ have a common non-singular p-adic zero for all p. This can for example be found in Birch's work [1], and applies to our case, since $\mathfrak{S}$ is absolutely convergent by Lemma 5.5.

Our singular integral can be treated in the very same way as in Schmidt's work [6]. First of all we know that $\tilde{J} > 0$, if $\dim V(0) = n_1 + n_2 - R$ and if the $F_i(\mathbf{x}; \mathbf{y})$ have a non-singular real zero in $\mathcal{B}_1 \times \mathcal{B}_2$. This is just Lemma 2 from Schmidt's paper [6]. Furthermore, we have shown in the proof of Lemma 5.6 that we have

$$|I(\boldsymbol{\beta})| \ll \min(1, \max_i |\beta_i|^{-R-1}),$$

which enables us to apply section 11 of [6]. This implies that the limit

$$\tilde{J} = \lim_{T \to \infty} \tilde{J}_T$$

exists and equals $\tilde{J} = J$. This proves our main theorem. □

References

[1] B. J. Birch, *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961), 245–263.
[2] H. Davenport, *Cubic Forms in Thirty-Two Variables*, Phil. Trans. R. Soc. Lond. A **251** (1959), 193–232.
[3] H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities*, Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2005. With a foreword by R. C. Vaughan, D. R. Heath-Brown and D. E. Freeman, Edited and prepared for publication by T. D. Browning.
[4] J. Harris, *Algebraic Geometry, A First Course*, Springer (1993).
[5] M. Robbiani, *On the number of rational points of bounded height on smooth bilinear hypersurfaces in biprojective space*, J. London Math. Soc. **63** (2001), 33–51.
[6] W. M. Schmidt, *Simultaneous rational zeros of quadratic forms*, Seminar Delange-Pisot-Poitou 1981. Progress in Math. Vol 22 (1982), 281–307.
[7] W. M. Schmidt, *The density of integer points on homogeneous varieties*, Acta Math. **154** (1985), no. 3-4 243–296.
[8] C. V. Spencer, *The Manin conjecture for $x_0y_0 + \ldots + x_sy_s = 0$*, J. Number Theory **129** (2009), no. 6, 1505–1521.
[9] K. van Valckenborgh, *Squareful numbers in hyperplanes*, arXiv 2011, 1001.3296v3.
[10] R. C. Vaughan, *The Hardy-Littlewood method*, volume 125 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, second edition, 1997.

School of Mathematics, University of Bristol, University Walk, Clifton, Bristol BS8 1TW, United Kingdom

E-mail address: `maxds@bristol.ac.uk`